Learning mechanism used when categorizing mathematical information surrounded by perceptual features

Abigail K. Mills

The Ohio State University

Author Note

I received the Office of Undergraduate Research & Creative Inquiry Summer Research Fellowship, which provided me with funding during the summer of 2017, as well as the OSU Arts & Sciences Undergraduate Research Scholarship, which funded this research during the academic term. Special thanks to Dr. Tasha Posid and Dr. Vladimir Sloutsky for their collaboration on this project. I would also like to thank my thesis committee members: Dr. Vladimir Sloutsky, Dr. Tasha Posid, Dr. Christopher Hoch, and Dr. Courtney McGinty.

Abstract

Much research indicates that fractions and proportions are difficult concepts to grasp. Is it possible to capitalize on a non-mathematical skill that we already possess to facilitate an understanding of difficult math concepts? We tested undergraduate students on their mathematical knowledge, including fractions, and followed this with a categorization task in which they learned a fraction concept. Our results show that adults easily learned a novel fraction rule across a variety of presentation conditions within our categorization task. However, accuracy was lower and reaction time was slower in conditions where participants were presented with additional and extraneous perceptual features, which presumably distracted them from the critical numerical information. This data mirrors that found with children, suggesting the critical role of visual attention and task demands rather than a developmental shift per se. By inducing adults to think like children with the introduction of challenging task demands, we can begin to understand the mechanism underlying children's learning, which will allow for better development of learning materials. Using a well-mastered skill (categorization) to learn a difficult math concept (fractions) without the presence of distracting perceptual information intruding on learning and transfer is a novel finding and may be a unique strategy for teaching other difficult concepts both inside and outside of formal education.

Keywords: Categorization, category learning, fractions

Learning mechanism used when categorizing mathematical information surrounded by perceptual features

Introduction

In most elementary classrooms in the United States, one can find an extensive number of colorful, visually appealing posters and learning tools, such as plastic blocks used for basic arithmetic or images of sliced pizza to teach fractions, that are used to keep the students' attention and facilitate learning (National Council of Teachers of Mathematics, 2000; Petersen & McNeil, 2012; Van de Walle, 2007). In fact, teachers report using perceptually appealing stimuli as a method to keep students focused on the task at hand (Petersen & McNeil, 2012). Despite the prevalence of these visually interesting education materials, there is a growing body of research to suggest that this is not always the best method of teaching (e.g., Petersen and McNeil, 2012). As such, it is important to understand the effects of these salient and perceptually rich educational materials in classrooms on children's mathematical learning in order to enhance the learning outcomes of this country's children. This is particularly important given that children in the United States do not demonstrate age-appropriate math skills and fall below their peers in other countries (National Center for Education Statistics, 2010; Siegler et al., 2012; also see Hurst & Cordes, 2016). Therefore, improving students' math knowledge and reasoning ability about fractions and proportions early in a child's education is important and understanding what instructional format may best lead to both the learning and transfer of difficult math concepts should be a fundamental component of instruction and curriculum development.

Perceptual variability and math learning

Many studies have revealed that having visually appealing, perceptually rich items to keep children's attention is not universally best for their learning. For example, Kaminski and

Sloutsky (2013) studied children's abilities to read bar graphs in the presence and absence of extraneous features in the bars, such as having flowers as shading versus a solid (monochromatic) black bar. Their results revealed that the presence of the extraneous features hindered children's ability to read bar graphs (Kaminski & Sloutsky, 2013). They found that when the number of objects contained in the bars matched the y-axis value (e.g., five flowers corresponding to an answer of five from the y-axis), children could correctly read and interpret the graph, but often gave incorrect answers if the number of items in the bars didn't match the y-value of the bar (e.g., three flowers corresponding to an answer of five from the y-axis; Kaminski & Sloutsky, 2013). Individuals who learned graph reading without extraneous objects were much more accurate in their graph reading upon test, thus allowing researchers to conclude the presence of extraneous perceptual features hindered learning, with more severe effects on children of a younger age (Kaminski & Sloutsky, 2013). This study indicates that when learning a new basic mathematical concept, perceptual features inhibit children's ability to learn the appropriate rule, while giving less visually appealing stimuli during training can actually foster better learning of the target mathematical skill.

In another study, Kaminski and Sloutsky (2012) showed that children learning with generic instantiations outperform their peers who learn with concrete stimuli. For example, children in the Generic condition were presented with the fraction "2/5" in its basic symbolic form and shown simple black and white circles to denote the numerator and denominator, while children in the Concrete condition were shown colorful flowers along with the "2/5" fraction (Kaminski & Sloutsky, 2012). Learning scores for children in the Generic condition were significantly higher than those in the Concrete condition (Kaminski & Sloutsky, 2012). When generalization of the fraction rules was tested in a transfer task of non-trained proportional reasoning, children in the Generic condition still outperformed their counterparts in the Concrete condition immediately after learning (Kaminski & Sloutsky, 2012). This study indicates that the concreteness of learning materials hinders the ability of a child to acquire new math knowledge. It suggests that the more concrete something is, the more likely a child is to be distracted by the extraneous information present in the concrete features that already have meaning on their own.

In a similar vein, a study by Koedinger, Alibali, and Nathan (2008) presented inexperienced and experienced algebra students with mathematical story problems of varying complexity and equations that corresponded to the stories. The results revealed that both inexperienced and experienced students were more successful on complex problems when using a more formal, abstract strategy, but that this also held true for experienced students on the simpler story problems, while inexperienced students performed better when using more informal, concrete strategies to solve simpler problems (Koedinger et al., 2008). This study suggests that, especially for experienced students, abstract, formal strategies are more beneficial when solving mathematical story problems than grounded, concrete strategies.

Developing a strategy to improve math learning

The ability to abstract the numerical information from a set, independent of its perceptual attributes (e.g., color, shape, texture, size…) is critical for a true understanding of number (Posid & Cordes, 2014a, for a review). Numerical abstraction has been observed in children as young as 3-4 years of age. For example, Posid and Cordes (2014a) asked children between the ages of three and six years old to determine which of two sets of objects contained a target number (e.g., 6 or 12), and found that children were more accurate on sets that were homogenous (i.e., all the same) in make-up rather than heterogeneous (i.e., all different; also see infancy research: Mix, 1999, 2008a, 2008b). Therefore, perhaps utilizing this abstraction ability could facilitate knowledge of a more challenging topic; that is, perhaps the use of a broader, yet related, ability – that is, categorization – can facilitate an understanding of difficult math concepts, such as proportions or fractions, particularly in the face of extraneous perceptual features that might otherwise hinder young children's learning.

More generally, a human's cognitive abilities are enhanced by categorization, the "process of grouping distinguishable entities into equivalence classes," thereby allowing us to generalize, communicate, and make inferences about a wide range of topics (Rivera & Sloutsky, 2015). Like abstraction, the ability to categorize (shapes and other non-numerical features) has also been demonstrated in preschool-age children (Deng & Sloutsky, 2015). Research has shown that adults are more advanced than children in attention optimization, which is the ability to focus on a feature that distinguishes one category from another while ignoring other irrelevant features (Hoffman & Rehder, 2010). That being said, children are still able to categorize with a certain level of accuracy even compared to proficient adults. Therefore, perhaps categorization – a skill that young children already possess – can facilitate an understanding of novel or difficult math concepts (in our case, fractions), as both skills rely on the use of abstraction. The current study explores that possibility. One point of concern is the fact that young children simply find an increasing level of perceptual features to be more distracting while categorizing a novel math concept and may be less able to attend to one single feature (Deng & Sloutsky, 2015, 2016; Plebanek & Sloutsky, 2017; Rivera & Sloutsky, 2015; Sloutsky, 2010; Sloutsky & Fisher, 2004, Sloutsky, Kloos, & Fisher, 2007; Hoffman & Rehder, 2010). The current study explores the relative impact of perceptual features on one's ability to categorize novel math information.

Current Study

Fractions are an integral part of our daily lives, from cooking to monetary calculations, so it is imperative to understand the mechanism children use when learning about fractions so that their educational curriculum can appropriately meet their needs. The National Mathematics Advisory Panel's 2008 report states that students in the United States perform at a less-satisfactory level of mathematical proficiency than that of their international peers. This is critical to address because competence with fractions is a significant factor in the successful learning of algebra, as well as a predictor of mathematical performance in high school, college, and careers in science, technology, engineering, and mathematics (STEM; National Mathematics Advisory Panel, 2008; Bailey, Hoard, Nugent, & Geary, 2012). In addition, students who aspire to enroll in advanced math and science classes in high school and begin to study algebra while in middle school have a distinct advantage over those who wait to take algebra until later and students who start at a lower level of ability tend to stay behind their peers who start with a greater understanding of fractions and math (Department of Education, 1997). Thus, the purpose of this study is to capitalize on individuals' strong abilities to categorize newly learned information and apply this ability to learning a fraction concept.

The goal of this study was to understand how our advanced adult categorization ability might aid children's less advanced understanding (or learning) of difficult math concepts, specifically proportions and fractions. We utilized adult participants to explore this question, in an effort to inform our work on children's understanding and learning of this phenomenon. Prior studies have found that perceptual features often distract from the deterministic features being taught or learned, thereby inhibiting learning (e.g., Kaminski & Sloutsky, 2013; Posid & Cordes, 2014a; Posid & Sloutsky, 2016). Does this also apply to one's ability to categorize information that is presented mathematically? I hypothesize that adults will be able to use their advanced numerical abstraction ability to make mathematical inferences when presented with novel categories that are mathematical in nature. Specifically, I predict that adults will be able to categorize mathematical information, while ignoring extraneous perceptual features, in an otherwise difficult task, but that the nature of the stimuli used will impact their learning and overall accuracy.

Methods

Participants

This study used undergraduate students at The Ohio State University participating in the Research Experience Program (REP) who received class credit for their involvement in the experiment. As detailed in Table 1, there were 97 participants overall and they were randomly assigned to one of four testing conditions. Written informed consent was obtained from all participants before they began the study.

Materials

Fraction and Math Battery

Individuals were first tested in a fraction battery consisting of three fraction-knowledge tasks that required participants to make judgements on a variety of symbolic or visual information (Hurst & Cordes, 2016; Posid & Sloutsky, 2016, 2017; Figure 1). Three tasks made up our Fraction Battery and were selected based on the Common Core's use of similar materials (Figure 1). The first task was an Ordinal Task, which asked participants to determine which of two options was numerically larger. There were three blocks, each containing 32 test items. The first block had whole number comparisons (1 versus 2). The second block had visual fraction pairs (black-and-white circles). The third block utilized symbolic fraction sets (1/5 versus 4/5).

The second task was a Matching Task that required participants to match a visual fraction, given as part of a shaded circle, to a symbolic fraction, or vice versa. The Matching Task was made up of three blocks, each containing 12 trials. The third task was an Addition and Subtraction Task where participants were asked to add symbolic and visual fractions (e.g., symbolic: 2/6 + 3/6, visual: they would be given two circles with six equal parts, one with two parts shaded and the other with three parts shaded). Four multiple choice answers were provided in both the Matching and Addition and Subtraction Tasks because of the potential difficulty level and all fractions were < 1, using no improper fractions throughout the tasks (per Hurst & Cordes, 2016; Posid & Sloutsky, 2016, 2017). Accuracy and reaction times on each of these tasks were combined into a Fraction Battery Accuracy score and Fraction Battery Reaction Time score for each individual participant, indicating their prior fraction knowledge.

Additionally, participants were presented with a speeded arithmetic sheet consisting of 160 basic arithmetic problems and were told to complete as many problems as possible in three minutes. Unlike the fraction task, this task was used to assess general math knowledge and fluency (as a sum score of correctly answered questions).

Categorization Task

Participants were randomly assigned to one of four conditions (see Figure 2): Full Arrays, Full Lines, Minimum Arrays, and Minimum Lines. The "full" distinction meant an entire character (designated as an "alien") was shown, including perceptual features (e.g., head, antenna, legs…), while the "minimum" condition presented just the "alien's" stomach and the deterministic feature (Figures 2 and 3; the new category rule; in this study, the alien's belly buttons; per Deng & Sloutsky, 2013, 2015; Posid & Sloutsky, 2016). In the Lines condition, the deterministic features were presented in organized lines while the Arrays condition had deterministic features randomly placed in an “array,” (see Figure 2). Lines vs. Arrays were included in the present study because one could argue that having the to-be-learned ratio of buttons in lines could be memorized based on “shape” or “layout” of the buttons (for a review, see Cantrell & Smith, 2013; Cordes & Brannon, 2009; Posid & Cordes, 2014a). Thus, the Arrays conditions control for approximate area across presentations, discouraging the use of these perceptually overlapping features as cues for learning.

Across conditions, participants were presented with two novel categories and were asked to determine which category a specific exemplar should be placed into. The categories were presented as two types of “aliens”. The first category was labeled as a “Flurp” (where the deterministic features, or the “belly buttons,” were presented as a 1/3 fraction) and the second category was labeled as a “Jalet” (“belly buttons” presented as a 1/2 fraction; see Figure 3). Participants were initially explicitly told that Flurps have buttons in a 1/3 ratio (1/3 or 2/6 fraction), while Jalets have buttons in a 1/2 ratio (1/2 or 3/6 fraction), with no mention of the other aspects of appearance associated with each.

In training, participants saw stimuli from our High Match Trial Type (Figure 4), meaning the images they were shown had the fraction-ratio corresponding to the respective category and the perceptual features were representative of the character’s appearance (i.e., 4 out of the 6 non-deterministic perceptual features were representative of the prototypical category; per Deng & Sloutsky, 2015). Participants had to correctly categorize nine “aliens” in a row or get nine out of ten categories correct before they were permitted to move on to the test phase. If this learning criterion was not reached, participants completed up to 60 training trials.

The test trials contained five trial types, which manipulated either the perceptual (P) or deterministic (D) features of the stimuli to be categorized. The five trial types were all modeled after the stimuli used in Deng & Sloutsky, 2013, 2015; Posid & Sloutsky, 2016; see Figure 4. As discussed previously, High Match stimuli included perceptual and deterministic features following the same rules from training (P_{Flurp}, D_{Flurp} and P_{Jalet}, D_{Jalet}; e.g., the alien perceptually looked like a Flurp and had Flurp deterministic buttons) and were included to determine how well participants had learned the two categories and whether they recognized the old items from training. Switch trials occurred when the perceptual features for one alien were presented with the deterministic features of the other (P_{Flurp}, D_{Jalet} and P_{Jalet}, D_{Flurp}; e.g., the alien perceptually looked like a Jalet but had Flurp deterministic buttons) and allowed us to determine if individuals were relying on the perceptual features, or overall similarity of the "aliens," or the actual rule as determined by the "belly buttons," the deterministic features. All-New-P trials included perceptual features that were completely different from those given in training (P_{New}, D_{Flurp} and P_{New}, D_{Jalet}; e.g., the alien perceptually looked brand new, but had Flurp deterministic features) to determine if participants can truly perform rule-based categorization, since they had no recognizable perceptual features to rely upon. One-New-P trials included one new perceptual feature (P_{OneNew}, D_{Flurp} and P_{OneNew}, D_{Jalet}; e.g., the alien looked mostly like a Flurp with a new feature, and had Flurp deterministic buttons) and were included to assess categorization abilities when a new perceptual feature is introduced, also judging if all perceptual features were encoded during training and test (similar to High Match trials). New-D trials gave a new deterministic feature not corresponding to a one-half or one-third ratio (P_{Flurp}, D_{New} and P_{Jalet}, D_{New}; e.g., the alien perceptually looked like a Flurp, but had new deterministic buttons not belonging to either learned category) and were used to assess how thoroughly participants encoded the deterministic rule. That is, if adults were overly focused on the deterministic features, they should respond at-chance for these questions. However, if adults, like children, distributed their attention during training trials, they should be able to correctly categorize the “alien” by its perceptual features (per Deng & Sloutsky, 2015).

Participants learned about two novel categories: “Flurps” had a one-third ratio of buttons and “Jalets” had a one-half ratio of buttons. In training, participants viewed stimuli identified as Flurps with arrangements of either 1/3 or 2/6. In test, exemplars in this category could have a 1/3 or 2/6 ratio of buttons, but they could also have a 3/9 ratio of buttons (called “New Exemplar Trials”). Similarly, Jalets were first shown as having arrangements of either 1/2 or 3/6 in training, and then in test, an additional ratio of 5/10 was shown in some cases. These New Exemplar trials allowed us to make conclusions about individuals’ abilities to generalize the fraction rule to novel situations.

All participants completed a total of 60 test trials (5 trial types X 3 fractions X 2 categories – repeated twice). Each exemplar was presented for 500 ms and participants clicked on one of two boxes (which remained on the screen even after the exemplar disappeared) to indicate their answer.

Procedure

Adult participants were tested at The Ohio State University in a quiet room in the Cognitive Development Lab. A female experimenter provided a laptop with a 13-inch screen for the participant to use. She set up each program for the participant and gave brief instructions for each task before allowing the participant to complete the task on their own.

Fraction and Math Battery

Each session began with the Fraction Battery, consisting of three fraction tasks to test prior fraction knowledge, followed by the categorization task, and finished with a timed math sheet. The fraction tasks consisted of an Ordinal Comparison Task, a Matching Task (fractions presented as numbers or black-and-white circles), and an Addition and Subtraction Task (symbols or black-and-white circles; per Hurst & Cordes, 2016; Posid and Sloutsky, 2016, 2017). Participants were given a timed math sheet after the categorization task to give us more information on their general math ability.

Categorization Task

Following the fraction tasks, the participant completed our critical categorization task. There were four versions of the game (2x2 design; see Figure 2): Full Lines, Full Arrays, Minimum Lines, and Minimum Arrays. The procedure across each Condition was identical and they only differed in perceptual make-up, as described previously.

In the first half of the task, training trials were given with feedback until a specific learning criterion (nine consecutive correct trials or nine out of ten correct trials) was reached; otherwise, the participants went through 60 trials before moving on to the test phase (Deng & Sloutsky, 2013, 2015; Posid & Sloutsky, 2016).

Following training, the test phase consisted of 60 trials where participants again had to judge whether the creatures presented were “Flurps” or “Jalets,” based on their deterministic features, as described in the Materials.

All programs were run and displayed on a MacBook laptop using RealBasic, which recorded participants’ accuracy and reaction time.

Data Analysis

To address our research questions, a series of planned analyses was run. First, a 2x2 mixed measures ANOVA was run (Condition: Full Arrays, Full Lines, Minimum Arrays, and Minimum Lines on Accuracy) to address (a) whether adults can learn a fraction rule in a categorization task and (b) under what perceptual and task conditions participants most benefit.

Secondary regression analyses were also run in order to examine the predictors of learning and retention in the categorization task, measuring the dependent variables of Trials to Learn, Categorization Task Accuracy, New Exemplar Accuracy, and Categorization Task Reaction Time, described later.

Results

Categorization Task

Adults successfully learned a fraction rule in our categorization task. When comparing their average accuracy to chance-level (.50), the mean score was 0.84 for both practice ($t(96) = 24.5$, $p < 0.001$, Cohen's $d = 5.0$) and test ($t(96) = 30.5$, $p < 0.001$, Cohen's $d = 6.2$). This above-chance performance on overall test accuracy held across the four conditions as well (Full Arrays: $M = 0.76$; $t(23) = 8.3$, $p < 0.001$, Cohen's $d = 3.5$; Minimum Arrays: $M = 0.84$; $t(23) = 19.3$, $p < 0.001$, Cohen's $d = 8.0$; Full Lines: $M = 0.89$; $t(26) = 22.4$, $p < 0.001$, Cohen's $d = 8.8$; Minimum Lines: $M = 0.88$; $t(21) = 59.2$, $p < 0.001$, Cohen's $d = 25.8$; see Figures 5 and 6). Individual differences in effect sizes across the conditions indicated individual and condition variability and warranted further analyses.

We further explored participants' accuracy through a repeated measures ANOVA. We examined whether accuracy in the Categorization Task interacted with testing condition. A repeated measures ANOVA was run without including the New-D[1] type [4 (Trial Type) X 2

[1] We first ran a 5 (Trial Type) X 2 (Full vs. Minimum) X 2 (Lines vs. Arrays) repeated measures ANOVA, with all values shown in Table 2. This first analysis revealed a main effect of Trial Type ($F(4, 372) = 272.6$, $p < 0.001$, $\eta_p^2 = 0.746$), with equal performance across High Match, One-New-P, All-New-P, and Switch trials (all *p's* < 0.001, Cohen's $d > 5.0$). Accuracy was at-chance for New-D ($M = 0.41$, $t(96) = -4.3$, $p < 0.001$, Cohen's $d = 0.89$). Prior studies have indicated that an at-chance performance on New-D trials means that participants are not actually relying on perceptual features during this task and are instead fully relying on deterministic features, which makes sense because we directed participants' attention to the deterministic features during training (Deng & Sloutsky, 2016). Therefore, if adults were looking at perceptual

(Full vs. Minimum) X 2 (Lines vs. Arrays)]. There was no main effect of Trial Type ($p > 0.80$). Critically, results revealed a main effect of Full vs. Minimum ($F(1{,}93) = 6.2$, $p = 0.015$, $\eta_p^2 = 0.062$), such that adults were more accurate on Minimum trials ($M = 0.94$) than Full trials ($M = 0.88$; see Figures 5 and 6). Unsurprisingly, there was also a main effect of Lines vs. Arrays ($F(1{,}93) = 15.6$, $p < 0.001$, $\eta_p^2 = 0.144$), such that participants were more accurate on Lines ($M = 0.96$) than Arrays ($M = 0.86$; see Figures 5 and 6). There was a marginal interaction between Full vs. Minimum and Lines vs. Arrays ($p = 0.06$; see Table 3), indicating an advantage for accuracy in the Minimum over Full conditions more so in the Arrays condition (Figure 5). There were no other main effects or interactions (all p's > 0.20).

We also examined participants' accuracy on our New Exemplar trials, that is, those ratios (3/9 or 5/10) that participants did not explicitly learn in training but were presented with during test trials. A 2 (Full vs. Minimum) X 2 (Lines vs. Arrays) univariate ANOVA with New Ratio Accuracy as the dependent variable was run and revealed no interaction ($p > 0.1$) between Full vs. Minimum and Lines vs. Arrays. However, there was a significant main effect of Full vs. Minimum ($F(1, 93) = 4.5$, $p = 0.036$, $\eta_p^2 = 0.046$) and a significant main effect of Lines vs. Arrays ($F(1, 93) = 13.4$, $p < 0.001$, $\eta_p^2 = 0.126$). Following the pattern that seems to have emerged with general accuracy, participants were better at Minimum ($M = 0.90$) than Full ($M = 0.83$) and better at Lines ($M = 0.93$) than Arrays ($M = 0.80$).

---

features – as children often do – adults in the present study should be responding above chance because the perceptual features correspond to the appropriate category (that is, you could use the perceptual OR deterministic features on all other trial types). Thus, we can conclude that adults used deterministic features and accurately learned the fraction rule in our task, and were simply answering at random when the deterministic features did not follow the learned rule (Deng & Sloutsky, 2016). As such, these results don't require any further analyses because they offer no new information. The repeated measures ANOVA also revealed an interaction between Full vs. Minimum and Trial Type ($F(4, 372) = 7.3$, $p < 0.001$, $\eta_p^2 = 0.073$), although this likely also had to do with the inclusion of New-D trials in our analyses.

Because adults performed so accurately in the categorization task, participants' reaction times were also assessed on the Categorization Task.[2] A 4 (Trial Type) X 2 (Full vs. Minimum) X 2 (Lines vs. Arrays) repeated measures ANOVA was run examining participants' average reaction time during test. A significant main effect of Trial Type ($F(3, 276) = 3.0$, $p = 0.032$, $\eta_p^2 = 0.031$) emerged even without the inclusion of New-D trials. This was largely due to the fact that participants were inexplicably slower on One-New-P trials (p's < 0.07) than on other trial types. There was a significant main effect of Lines vs. Arrays ($F(1, 92) = 42.7$, $p < 0.001$, $\eta_p^2 = 0.317$) that arose because, unsurprisingly, participants were faster at Lines ($M = 0.83$ s) than Arrays ($M = 1.2$ s). Critically, there was a significant interaction between Full vs. Minimum and Lines vs. Arrays ($F(1, 92) = 5.4$, $p = 0.022$, $\eta_p^2 = 0.056$). It appears that participants were slower at identifying the correct option with Arrays (Full $M = 1.1$ s; Minimum $M = 1.3$ s) than with Lines (Full $M = 0.85$ s; Minimum $M = 0.81$ s), such that the magnitude of the difference between Lines and Arrays in the Full Condition was larger than the magnitude of the difference between Lines and Arrays in the Minimal Condition.

Results also revealed an interaction between Trial Type and Lines vs. Arrays ($F(3, 276) = 3.4$, $p = 0.019$, $\eta_p^2 = 0.035$). Participants were slower when viewing Arrays across all four trial types ($M = 1.2$ s for all trial types) compared to Lines ($M_{average} = 0.83$ s; All-New-P: $M = 0.76$ s; High Match: $M = 0.87$ s; One-New-P: $M = 0.89$ s; Switch: $M = 0.81$ s); however, there was much more variability in reaction time across the Lines conditions than there was in the Arrays

[2] Reaction time is used as a dependent variable here because it shows "the speed at which perceptual information is processed," which is often used in developmental research because "a systematic increase in processing speed has been proposed to underlie cognitive development" (Duan, Shi, & Zhou, 2010, Ferguson & Bowey, 2005, Hale, 1990, Kail, 2000). This variable allows us to understand more about adults' thought processes, rather than just the observed skills and knowledge associated with reported accuracy. For instance, a slower reaction time would indicate less confidence in answering in that condition than in others, potentially hesitating to break a rule or the like (Posid & Sloutsky, 2016).

conditions. A 3-way interaction between Trial Type, Minimum vs. Full, and Lines vs. Arrays was also found ($F(3, 276) = 3.0$, $p = 0.031$, $\eta_p^2 = 0.032$; see Table 5). These results suggest that, because adults have been found to pay attention to deterministic features over perceptual features, they were likely using some sort of strategy (e.g., counting; see Polinsky, Posid, & Sloutsky, under revision; Posid & Cordes, 2017) to figure out the ratio of buttons, which was not necessary in the Lines conditions.

We were also interested in participants' reaction time on New Exemplar trials, that is, on trials in which the ratio was the same as the practiced examples but that had not been seen during training. A univariate ANOVA revealed a main effect of Full vs. Minimum ($F(1, 92) = 8.8$, $p = 0.004$, $\eta_p^2 = 0.087$), with individuals being faster with Full ($M = 1.0$ s) than they were with Minimum ($M = 1.3$ s) presentations. There was also a main effect of Lines vs. Arrays ($F(1, 92) = 46.7$, $p < 0.001$, $\eta_p^2 = 0.337$) that, consistent with the other Reaction Time findings, shows adults are faster when deterministic features are displayed as Lines ($M = 0.86$ s) compared to Arrays ($M = 1.5$ s). Last, there was a significant interaction between Full vs. Minimum and Lines vs. Arrays ($F(1, 92) = 5.1$, $p = 0.026$, $\eta_p^2 = 0.053$) that shows participants are generally slowest with Minimum Arrays ($M = 1.7$ s) when compared to the other three conditions: Minimum Lines ($M = 0.90$ s), Full Arrays ($M = 1.2$ s), and Full Lines ($M = 0.83$ s). Although the Minimum stimuli were overall more helpful for adults in our study, when they were presented with the difficult arrays, they responded more quickly with the Full stimuli presumably because they were trying to use the perceptual features to answer efficiently in the more difficult task.

Predictors of Learning

In addition to examining participants' performance on the Categorization Task, we also examined what factors accounted for the individual variability observed in this task. We ran a series of linear regressions to examine the effects of demographic variables, fraction and math battery performance, and condition on our dependent variables of interest (Trials to Learn, Categorization Task Accuracy, New Exemplar Accuracy, Categorization Task Reaction Time; see Figures 7-10).

The first regression included Trials to Learn as the dependent variable and the predictors were Year in School, Age, Gender, Full vs. Minimum, Lines vs. Arrays, Fraction Battery Accuracy, Math Accuracy, and Fraction Battery Reaction Time. The significant predictors of Trials to Learn were Age ($\beta = 0.289$, $p = 0.004$), Lines vs. Arrays ($\beta = -0.393$, $p < 0.001$; with faster learning in Lines than Arrays), and Fraction Battery Accuracy ($\beta = -0.301$, $p = 0.003$), while all others had $p > 0.1$; Model: $R^2 = 0.469$, $p < 0.001$), as shown in Figure 7. A secondary linear regression model included the accuracies of the three individual Fraction Battery tasks as predictors, rather than the Fraction Battery as a composite score. Ordinal Task Accuracy ($\beta = -0.282$, $p = 0.004$) was a significant predictor of Trials to Learn, as was Addition and Subtraction Task Accuracy ($\beta = -0.187$, $p = 0.031$; Age and Lines vs. Arrays remained significant: *p's* < 0.001). It is possible that adults are more adept at recognizing whether one-half or one-third is larger in the categorization task, which helps them distinguish between the two categories. If this is true, recalling a rote fact is beneficial for learning in the categorization task, indicated by the significance of the Ordinal Task, while ability in the Addition and Subtraction Task is an indicator of general fraction ability that helps participants better learn the fraction rule in fewer trials. The other predictors that were measured had $p > 0.1$ (Model: $R^2 = 0.499$, $p < 0.001$). Additionally, we ran a hierarchical regression (Model 1: Lines vs. Arrays; Model 2: Lines vs. Arrays, Fraction Battery Accuracy; Model 3: Lines vs. Arrays, Fraction Battery Accuracy, Age) to examine the relative contribution of each of our significant variables on our outcome variable.

Lines vs. Arrays alone accounted for 22.8% of the variance in our model ($p < 0.001$). The inclusion of Fraction Battery Accuracy accounted for an additional 10.4% of the variance in our model (both p's < 0.001). Finally, the inclusion of Age accounted for an additional 9.0% of the variance in our model, above and beyond that of Fraction Battery Accuracy (all p's < 0.001). This suggests that, although math ability and age tend to be correlated in other studies examining these variables, they uniquely and individually contributed to the variance observed in our outcome variable.

Accuracy on the test trials of our categorization task was significantly predicted by Age ($\beta = -0.341$, $p = 0.003$), Lines vs. Arrays ($\beta = 0.290$, $p = 0.003$), and Fraction Battery Accuracy ($\beta = 0.320$, $p = 0.005$), depicted in Figure 8. All other predictors had $p > 0.1$ (Model: $R^2 = 0.336$, $p < 0.001$). Since Fraction Battery Accuracy was a significant predictor of Test Accuracy, we ran a second linear regression and replaced Fraction Battery Accuracy with Ordinal Task Accuracy, Matching Task Accuracy, and Addition and Subtraction Task Accuracy as predictors. With the inclusion of the individual fraction tasks, we found that Math Accuracy ($\beta = -0.189$, $p = 0.093$, marginal) and the Addition and Subtraction Task ($\beta = 0.259$, $p = 0.009$) significantly predicted performance in the Categorization Task (Model: $R^2 = 0.344$, $p < 0.001$). It is possible that once the adults learned the fraction rule with assistance from rote fact recall, as mentioned for Trials to Learn, their ability to perform mental arithmetic quickly and with high accuracy helped them categorize the creatures more efficiently and with more accuracy. Again, Age ($\beta = -0.270$, $p = 0.015$) and Lines vs. Arrays ($\beta = 0.326$, $p = 0.001$) remained significant in this model (all other p's > 0.1). Also, we ran a hierarchical regression (Model 1: Lines vs. Arrays; Model 2: Lines vs. Arrays, Fraction Battery Accuracy; Model 3: Lines vs. Arrays, Fraction Battery Accuracy, Age) to examine the relative contribution of each of our significant variables on our outcome variable.

Lines vs. Arrays alone accounted for 14.4% of the variance in our model ($p < 0.001$). The inclusion of Fraction Battery Accuracy accounted for an additional 5.5% of the variance in our model (both p's < 0.001). Finally, the inclusion of Age accounted for an additional 5.3% of the variance in our model, above and beyond that of Fraction Battery Accuracy (all p's < 0.001). This suggests that, although math ability and age tend to be correlated in other studies examining these variables, they uniquely and individually contributed to the variance observed in our outcome variable. We additionally ran linear regressions for each trial type's accuracy with the predictors Age, Condition, Fraction Battery Accuracy, Fraction Battery Reaction Time, and Math Accuracy, which can be found in Table 6.

Pictured in Figure 9, we examined New Exemplar Accuracy as the dependent variable in another linear regression with the predictors Year in School, Age, Gender, Full vs. Minimum, Lines vs. Arrays, Math Accuracy, Fraction Battery Accuracy, and Fraction Battery Reaction Time. Age ($\beta = -0.317$, $p = 0.007$) and Lines vs. Arrays ($\beta = 0.339$, $p = 0.001$) were significant predictors of New Exemplar Accuracy (all others $p > 0.1$; Model: $R^2 = 0.269$, $p = 0.001$). Fraction Battery Accuracy was not run with its three individual task accuracies in a secondary regression because it was not found to be significant as a composite score.

Finally, a linear regression with Reaction Time on the Test Trials of our categorization task as the dependent variable was run. The predictors included in this model were Year in School, Age, Gender, Full vs. Minimum, Lines vs. Arrays, Math Sheet Accuracy, Fraction Battery Accuracy, and Fraction Battery Reaction Time. Reaction time on the Categorization task was significantly predicted by Lines vs. Arrays ($\beta = -0.444$, $p < 0.001$) and Fraction Battery Reaction Time ($\beta = 0.329$, $p = 0.005$), shown in Figure 10. Fraction Battery Accuracy was not a significant predictor ($p > 0.1$; Model: $R^2 = 0.313$, $p < 0.001$), so a secondary linear regression was not necessary.

From these regressions, we generally learned that age, condition, particularly Lines vs. Arrays, and prior math and fraction knowledge mattered most when learning the fraction rule during training and in participants' accuracy during the test trials of the categorization task. As discussed earlier, the Fraction Battery was broken down into the individual tasks and it was discovered that the task predictors varied with the outcome variable. As such, the Ordinal Task was more important when learning the rule initially, with some significance found with the Addition and Subtraction Task, while the latter and the Math Sheet score (speeded arithmetic) were significant predictors of accuracy in the categorization task, revealing prior math knowledge and mental arithmetic as important for these tasks. Further implications are discussed next.

Discussion

The Ohio Department of Education has found that students are not performing at an age-appropriate level on standardized math tests, despite changes to the Core Curriculum. Much research reveals that individuals in the United States are significantly behind their peers in other countries in mathematics skills, affecting success in school and STEM fields (National Mathematics Advisory Panel, 2008; Bailey et al., 2012). Thus, the National Mathematics Advisory Panel (2008) has emphasized a need to gather additional research on general mathematics learning by providing insight into the learning mechanisms of students in the United States. Education and the development of a successful curriculum is based on the principle of adding new content to foundational knowledge and using previously acquired skills to ease the learning process during the acquisition of new knowledge, so it is imperative we understand the basics of children's learning so we can build upon it with more advanced knowledge. The present study capitalized on this concept and examined how adult participants might use a skill like categorization to learn novel fraction concepts. Our goals for this study were to (1) identify whether adults can learn and generalize a fraction rule using a categorization task and (2) understand the effects of perceptual features on learning and generalization.

Results from this study demonstrate several important findings. First, data replicate the finding that adults can learn a novel non-numerical category (e.g. shape, Deng & Sloutsky, 2015, 2016) when it's presented in the form of a categorization task (also see Posid & Sloutsky, 2016; Posid, Mills, & Sloutsky, in preparation). Even when participants aren't directly taught a fraction concept, something like a categorization task can help them learn, demonstrated by at-ceiling performance on all but New-D Trial Types. Although unsurprising that adults could learn a fraction rule quickly and reliably, these data provide further evidence that a categorization task may be an avenue by which individuals can learn about difficult concepts like fractions. By directing adults' attention to the deterministic features (belly buttons) in the task, adults were able to use this ability to categorize in order to learn novel fraction information, rather than via a more traditional measure as would be used in the classroom. Thus, while Deng & Sloutsky (2015, 2016) explored categorization with simple shapes, our study has extended these findings to demonstrate the benefit of categorization with numerical stimuli.

Results also indicate that Condition significantly affected both participants' learning and accuracy in the categorization task. Unsurprisingly, participants learned fastest when presented with Lines rather than Arrays. In line with previous research suggesting cumulative perceptual cues may lead to better discrimination, we found this to be true in our categorization task as well (Cordes & Brannon, 2009; Posid & Cordes, 2014a). In addition, adults may have been using an effortful strategy to figure out the ratio presented in the Arrays, which was not necessary in the Lines conditions. For example, some research suggests that increased reaction time (and/or increased fixations, as per eye-tracking studies) indicates the use of an effortful strategy, such as counting (e.g., Polinsky et. al., under revision; also see: Lipton & Spelke, 2005; Posid & Cordes, 2014b, 2017; Wylie, Jordan, & Mulhern, 2012). Thus, adults are usually skilled at directing their attention to relevant information quickly while ignoring irrelevant features, but they are best at doing so when the deterministic feature is presented in an organized, more easily understood format.

Importantly, participants performed most accurately in the Minimum condition compared to the Full Condition. These findings align with previous research demonstrating that extraneous information that is irrelevant to a distinguishing rule is distracting when individuals are trying to learn novel and/or task-relevant information. For instance, Kaminski and Sloutsky (2013) found that bar graphs filled with objects, like flowers, inhibited appropriation of the skill of graph reading when compared to that using solid black bars. Additionally, a study that used concrete stimuli, meaning objects that already have meaning like flowers, and generic stimuli, numeric symbols in this case, to teach fractions found better learning scores both initially and after a delay in children who learned with generic features instead of concrete ones (Kaminski & Sloutsky, 2012). Thus, the less prior meaning learning materials already contain and more visually simple they are, the less likely individuals are to be distracted in the learning process, resulting in better, more generalizable learning overall.

This “less is more” finding holds across development. As part of a larger study, the data presented here was compared with data using the same general methods with child participants to help us understand similarities in learning across development (Posid, et al., in preparation).

Generally, adults tend to do well with and without perceptual features because they are better at selectively attending to the deterministic features. However, in this experiment and in one part of the larger study using child participants, the stimuli viewed and categorized were shown briefly enough (500 ms) that task demands increased and forced adults to perform the task much like children might do, who would naturally find this type of task quite difficult, as compared to stimuli shown for 10 seconds in other portions where adults performed better than children. We believe the short exposure time increased the overall task difficulty and led adults to attempt to rely on both the perceptual and deterministic features (distributed attention) when making their decisions (Best et al., 2013; Deng & Sloutsky, 2016; Hoffman & Rehder, 2010, Sloutsky & Fisher, 2004; Sloutsky et al., 2007). These findings expand on the current literature to suggest that not only does the type of stimuli matter (full vs. minimum), but that task difficulty (i.e., task demands) rather than a developmental shift in one's ability to categorize and selectively attend to relevant information per se is important.

Although findings from this study generally support a "less is more" notion for categorization of novel math concepts, some work does indicate that this is not always the case. For example, a study by Posid and Sloutsky (2017) suggests that "less" is not always "more" when trying to teach children new mathematical concepts. In training first and second grade children on unlearned fraction concepts, the authors ran children in one of three training conditions: (1) Visual (Picture) Only, (2) Symbols Only, and (3) Visual + Symbols, across several fraction teaching sessions (Posid & Sloutsky, 2017). In general, the authors found that children performed least accurately following training in the Visual (Picture) Only condition, with highest success following training in the Visual + Symbols condition (Posid & Sloutsky, 2017). Thus, these findings led to the conclusion that children benefit from redundant perceptual information, such as when visual fractions were paired with the fraction symbols (Posid & Sloutsky, 2017). These results seem to concur with the findings in Petersen and McNeil (2012) that also found the presence of perceptual features can be beneficial for learning in *certain cases*. These results provide further support for the role of task demands in mathematical tasks, and attention should be paid to this when investigating the role of perceptual attributes on learning and attention.

The present study also investigated what variables may predict participants' learning and accuracy in our categorization task. When the predictors of learning were examined, there was a significant effect of age, which was potentially surprising given that children as young as 4 years of age can categorize proficiently. The effect of age likely arose from difference in experience between participants, such that individuals who have taken more math courses or who have more math-centered majors may be more fluent in their use of math concepts. Additionally, the prefrontal cortex, a brain region responsible for higher level thinking, is not fully mature in most individuals until 25 years of age, so the fact most of our participants were younger than this means the effect of age could have been explained by varied development in prefrontal cortex maturity. We suspect that, for example, given a sample of older adults (e.g., 40-45 years of age), these age effects would not be seen.

The significance of Condition also emerged above and beyond our other mathematic and demographic variables, as discussed above. However, there was a significant impact of prior math and fraction knowledge. This is interesting given that adults in our task should have been proficient in these basic fraction concepts. However, we also know that fractions and proportions, unlike other mathematics concepts, are notoriously difficult. For example, DeWolf, Bassok, and Holyoak (2015) found that in a complex analogical reasoning task, fraction notation had a disadvantage compared to decimal notation. An automatic response was not observed when comparing fractions, especially in cases where the larger denominator was part of the smaller fraction, because these values required more cognitive work than those where the larger denominator indicated the larger fraction (DeWolf et al., 2015). As the deterministic features (belly buttons) in our study are distinct entities and easy to divide into visual numerators and denominators, this allowed a more automatic response to take place because the buttons facilitate adults' innate preference for one-to-one mapping between number components and visual stimuli (DeWolf et al., 2015).

The finding that prior math knowledge does impact accuracy on our novel fraction task is interesting given that adults do have a well-developed and abstract concept of fractions. Reaction time and accuracy on the Fraction Battery were significant predictors of learning and generalization, indicating that the more familiarity and speed participants had with math and fractions benefitted their performance in a novel task. This is consistent with a finding from the Department of Education in 1997 that shows individuals who start with more capabilities stay ahead of their peers who begin with fewer skills in basic math (also see Kloosterman, 2010; Stigler, Givven, & Thompson, 2009; Stayflidou & Vosniadou, 2004).

The present study speaks more broadly to the possible mechanisms between the individual- and condition-based variability observed in our categorization task, even amongst advanced adults. This study – along with related work from our lab (Deng & Sloutsky, 2015; Plebanek & Sloutsky, 2017; Polinsky, et al., under revision; Posid & Sloutsky, 2016; Posid et al., in preparation) – suggests the role of attention optimization and selective attention (also see: Best, Yim, & Sloutsky, 2013; Posner & Rothbart, 2007; Rueda, Posner, & Rothbart, 2005). In Posid et al. (in preparation), this current experiment was compared to one in which stimuli were shown to adults for more than just a few seconds in test. Adults performed markedly better when there was not a time limit but performed more like children when the stimuli were shown only briefly (i.e., making the task more difficult; Posid et al., in preparation). In this sense, it is feasible that selective attention does develop with age, but that this cannot make up for the effects of challenging task demands. Thus, when adults' mental functioning is taxed to a high extent, as seen in this study with timed stimuli presentations, they categorize differently in the context of a math task than they would when the load on their working memory isn't as heavy. These effects of the timed stimuli on adults may mimic the effects of novel learning in children such that we can interpret the results to reveal the ultimate learning mechanism in a similar manner.

In sum, this study has yielded many of the results predicted at the outset. Specifically, we found that perceptual features are a hindrance to the learning of novel fraction information in our categorization task. This work expands on current work in our laboratory (Posid et al., in preparation) such that more difficult task demands led to the emergence of Condition differences, similar to Condition effects observed in children. From this, we can conclude that in situations with additional task difficulty, adults perform like children when learning fractions and our findings can be used to better strengthen current math curriculums to aid children in learning this crucial skill more efficiently.

References

Bailey, D. H., Hoard, M. K., Nugent, L., & Geary, D. C. (2012). Competence with fractions predicts gains in mathematics achievement. *Journal of Experimental Child Psychology*, *113*(3), 447–455. http://doi.org/10.1016/j.jecp.2012.06.004

Best, C.A., Yim, H., & Sloutsky, V.M. (2013). The cost of selective attention in category learning: Developmental differences between adults and infants. *Journal of Experimental Child Psychology, 116*, 105-119.

Cantrell, L., & Smith, L. B. (2013). Open questions and a proposal: A critical review of the evidence on infant numerical abilities. *Cognition, 128*(3), 331-352.

Cordes, S., & Brannon, E. M. (2009). The relative salience of discrete and continuous quantity in young infants. *Developmental Science, 12*(3), 453-463.

Deng, W., & Sloutsky, V. M. (2013). The role of linguistic labels in inductive generalization. *Journal of Experimental Child Psychology, 114,* 432-455.

Deng, W., & Sloutsky, V. M. (2015). The development of categorization: Effects of classification and inference training on category representation. *Developmental Psychology, 51,* 392-405.

Deng, W., & Sloutsky, V. M. (2016). Selective attention, diffused attention, and the development of categorization. *Clinical Psychology, 91,* 24-62.

Department of Education (1997). *Mathematics equals opportunity.* White Paper prepared for US Secretary of Education Richard W. Riley. Washington, DC.

DeWolf, M., Bassok, M., & Holyoak, K. J. (2015). Conceptual structure and the procedural affordances of rational numbers: Relational reasoning with fractions and decimals. *Journal of Experimental Psychology, 144*(1), 127-150.

Duan, X., Shi, J., & Zhou, D. (2010). Developmental changes in processing speed: Influence of accelerated education for gifted children. *Gifted Child Quarterly, 54*(2), 85-91.

Ferguson, A. N., & Bowey, J. A. (2005). Global processing speed as a mediator of developmental changes in children's auditory memory span. *Journal of Experimental Child Psychology, 91*, 89-112.

Hale, S. (1990). A global developmental trend in cognitive processing speed. *Child Development, 61*, 653-663.

Hoffman, A. B., & Rehder, B. (2010). The costs of supervised classification: The effect of learning task on conceptual flexibility. *Journal of Experimental Psychology: General*, *139*, 319–340.

Hurst, M., & Cordes, S. (2016). Rational-number comparison across notation: Fractions, decimals, and whole numbers. *Journal of Experimental Psychology: Human Perception and Performance*, 42(2), 281-293.

Kail, R. (2000). Speed of information processing: Developmental change and links to intelligence. *Journal of School Psychology, 38*, 51-61.

Kaminski, J. A., & Sloutsky, V. M. (2012). Children's acquisition of fraction knowledge from concrete versus generic instantiations. *Proceedings of the XXXIV Annual Conference of the Cognitive Science Society,* 1750-1755.

Kaminski, J. A., & Sloutsky, V. M. (2013). Extraneous perceptual information interferes with children's acquisition of mathematical knowledge. *Journal of Educational Psychology, 105*, 351-363.

Kloosterman, P. (2010). Mathematics skills of 17-year-olds in the United States: 1978 to 2004. *Journal for Research in Mathematics Education*, 20-51.

Koedinger, K. R., Alibali, M. W., & Nathan, M. J. (2008). Trade-offs between grounded and abstract representations: Evidence from algebra problem solving. *Cognitive Science, 32*, 366-397.

Lipton, J. S., & Spelke, E. S. (2005). Preschool children's mapping of number words to nonsymbolic numerosities. *Child development*, *76*(5), 978-988.

Mix, K. S. (1999). Similarity and numerical equivalence: Appearances count. *Cognitive Development, 14,* 269-297.

Mix, K. S. (2008a). Children's equivalence judgments: Cross-mapping effects. *Cognitive Development, 23,* 191-203.

Mix, K. S. (2008b). Surface similarity and label knowledge impact early numerical comparisons. *British Journal of Developmental Psychology, 26,* 13-32.

National Center for Education Statistics (2010). *The nation's report card: Mathematics 2009: National assessment of educational progress at grades 4 and 8*. (NCES 2010-451). Institute of Education Sciences, U.S. Department of Education, Washington, D. C.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

National Mathematics Advisory Panel (2008). The final report of the National Mathematics Advisory Panel. *Foundations for Success*, i-90.

Ohio Department of Education. Retrieved from: http://education.ohio.gov/Topics/Ohios-Learning-Standards/Mathematics.

Petersen, L. A., & McNeil, N. M. (2012). Perceptually rich manipulatives on preschoolers' counting performance: Established knowledge counts. *Child Development.* doi: 10.1111/cdev.12028

Plebanek, D. J., & Sloutsky, V. M. (2017). Costs of selective attention: When children notice what adults miss. *Psychological Science*, *28*, 723-732.

Polinsky, N., Posid, T., & Sloutsky, V. M. (under revision).

Posid, T., Mills, A. K., & Sloutsky, V. M. (in preparation).

Posid, T., & Cordes, S. (2014a). Verbal counting moderates perceptual biases found in children's cardinality judgments. *Journal of Cognition and Development*. doi: 10.1080/15248372.2014.934372

Posid, T., & Cordes, S. (2014b). The small-large divide: A case of incompatible numerical representations in infancy. In D. Geary, D. Berch, & K. Mann-Koepke (Eds.). *Evolutionary Origins and Early Development of Basic Number Processing*.

Posid, T., & Cordes, S. (2017). How high can you count? Probing the limits of young children's counting. *Developmental Psychology*.

Posid, T., & Sloutsky, V. M. (2016). *Kindergarteners and adults learn fraction rules in a categorization task*. Poster presented at the annual meeting of the Cognitive Science Society. Philadelphia, Pennsylvania.

Posid, T., & Sloutsky, V. M. (2017). *When less isn't more: A real-world fraction intervention study.* Abstract presented at the biennial meeting of the Society for Research in Child Development. Austin, Texas.

Posner, M.I., & Rothbart, M.K. (2007). Research on attention networks as a model for the integration of psychological science. *Annual Review of Psychology, 58*, 1-23.

Rivera, S., & Sloutsky, V. M. (2015). *Development of selective attention in category learning*. Poster presented at the annual meeting of the Cognitive Science Society. Pasadena, California.

Rueda, M.R., Posner, M.I., & Rothbart, M.K. (2005). The development of executive attention: Contributions to the emergence of self-regulation. *Developmental Neuropsychology, 28*, 573-594.

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., Susperreguy, M. I., & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science, 23*.

Sloutsky, V. M. (2010). From perceptual categories to concepts: What develops? *Cognitive Science*, *34*, 1244-1286.

Sloutsky, V. M., & Fisher, A. V. (2004). When development and learning decrease memory. *Psychological Science*, *15*, 553-558.

Sloutsky, V. M., Kloos, H., & Fisher, A. V. (2007). When looks are everything. *Psychological Science*, *18*, 179-185.

Stayflidou, S., & Vosniadou, S. (2004). The development of students' understanding of the numerical value of fractions. *Learning and instruction*, *14*(5), 503-518.

Stigler, J. W., Givvin, K. B., & Thompson, B. J. (2010). What community college developmental mathematics students understand about mathematics. *MathAMATYC Educator*, *1*(3), 4-16.

Van de Walle, J. A. (2007). *Elementary and middle school mathematics: Teaching developmentally*. New York: Pearson.

Wylie, J., Jordan, J. A., & Mulhern, G. (2012). Strategic development in exact calculation: Group and individual differences in four achievement subtypes. *Journal of experimental child psychology*, *113*(1), 112-130.

	Full Arrays	Minimum Arrays	Full Lines	Minimum Lines
Total Number	$N = 24$	$N = 24$	$N = 27$	$N = 22$
Gender	10 females, 14 males	13 females, 11 males	14 females 13 males	6 females 16 males
Age	$M = 19.3$ yrs, $SD = 3.3$ yrs	$M = 20.4$ yrs, $SD = 5.0$ yrs	$M = 19.2$ yrs, $SD = 1.1$ yrs	$M = 19.3$ yrs, $SD = 1.0$ yrs
Years in College	$M = 1.42$	$M = 1.42$	$M = 1.46$	$M = 1.64$

Table 1. Distribution of participants in each condition of the categorization task, including the number of participants, their self-identified genders, age in years, and number of years in college.

	High Match	One-New-P	All-New-P	Switch	New-D
Full Arrays	$p < 0.001$; $d = 3.7$	$p < 0.001$; $d = 3.2$	$p < 0.001$; $d = 2.8$	$p < 0.001$; $d = 2.4$	$p = 0.144$; $d = 0.63$
Minimum Arrays	$p < 0.001$; $d = 7.7$	$p < 0.001$; $d = 5.0$	$p < 0.001$; $d = 5.5$	$p < 0.001$; $d = 8.7$	$p < 0.001$; $d = 1.8$
Full Lines	$p < 0.001$; $d = 8.9$	$p < 0.001$; $d = 8.8$	$p < 0.001$; $d = 6.9$	$p < 0.001$; $d = 6.9$	$p = 0.646$; $d = 0.18$
Minimum Lines	$p < 0.001$; $d = 23.2$	$p < 0.001$; $d = 11.5$	$p < 0.001$; $d = 16.7$	$p < 0.001$; $d = 16.8$	$p = 0.42$; $d = 1.1$

Table 2. Significance and effect size (Cohen's *d*) for accuracy vs. chance across each trial type and condition in the test phase of the categorization task.

	High Match	One-New-P	All-New-P	Switch	New-D
Full vs. Minimum	$p = 0.058$; $\eta_p^2 = 0.038$	$p > 0.1$ $\eta_p^2 = 0.024$	$p = 0.024$; $\eta_p^2 = 0.053$	$p = 0.007$; $\eta_p^2 = 0.075$	$p = 0.013$; $\eta_p^2 = 0.065$
Lines vs. Arrays	$p = 0.006$; $\eta_p^2 = 0.079$	$p < 0.001$; $\eta_p^2 = 0.135$	$p = 0.002$; $\eta_p^2 = 0.098$	$p < 0.001$; $\eta_p^2 = 0.126$	$p = 0.032$; $\eta_p^2 = 0.048$

Table 3. Significance and effect size (η_p^2) for accuracy across Conditions (Full vs. Minimum or Lines vs. Arrays) of trial types in the categorization task. P-value indicates a difference between the Full vs. Minimum condition or a difference between the Lines vs. Arrays conditions within that particular trial type.

	High Match	One-New-P	All-New-P	Switch	New-D
Full vs. Minimum	$p > 0.2$ $\eta_p^2 = 0.015$	$p > 0.3$ $\eta_p^2 = 0.01$	$p > 0.2$ $\eta_p^2 = 0.016$	$p = 0.052$; $\eta_p^2 = 0.04$	$p > 0.1$ $\eta_p^2 = 0.029$
Lines vs. Arrays	$p < 0.001$; $\eta_p^2 = 0.262$	$p < 0.001$; $\eta_p^2 = 0.224$	$p < 0.001$; $\eta_p^2 = 0.29$	$p < 0.001$; $\eta_p^2 = 0.273$	$p > 0.2$ $\eta_p^2 = 0.014$

Table 4. Significance and effect size (η_p^2) for reaction time (seconds) across Conditions (Full vs. Minimum or Lines vs. Arrays) of trial types in the categorization task. P-value indicates a difference between the Full vs. Minimum condition or a difference between the Lines vs. Arrays conditions within that particular trial type.

	High Match	One-New-P	All-New-P	Switch
Minimum Arrays	1.2	1.3	1.3	1.4
Minimum Lines	0.85	0.88	0.74	0.76
Full Arrays	1.1	1.2	1.1	1.0
Full Lines	0.89	0.89	0.77	0.85

Table 5. Reaction times (in seconds) for each trial type (excluding New-D) across the four conditions, demonstrating a 3-way interaction between Trial Type, Full vs. Minimum, and Lines vs. Arrays in the categorization task.

	Age	Condition	Math Sheet	Fraction Battery Accuracy	Fraction Battery Reaction Time
High Match	$p < 0.001$ $\beta = -0.381$	$p = 0.006$ $\beta = 0.264$	p=0.810 $\beta = -0.029$	$p = 0.005$ $\beta = 0.303$	$p = 0.041$ $\beta = 0.232$
One-New-P	$p < 0.001$; $\beta = -0.370$	$p < 0.001$; $\beta = 0.366$	$p = 0.946$ $\beta = -0.008$	$p = 0.092$; $\beta = 0.174$	$p = 0.490$ β 0.076
All-New-P	$p = 0.027$; $\beta = -0.228$	$p < 0.001$; $\beta = 0.363$	p =0.662 $\beta = 0.054$	$p = 0.071$; $\beta = 0.196$	$p = 0.024$; $\beta = 0.265$
Switch	$p = 0.279$ $\beta = -0.113$	$p < 0.001$; $\beta = 0.411$	$p = 0.759$ $\beta = -0.039$	$p = 0.382$ $\beta = 0.097$	$p = 0.297$ $\beta = 0.124$
New-D	$p = 0.044$; $\beta = 0.225$	$p = 0.288$ $\beta = 0.111$	$p = 0.336$ $\beta = -0.130$	$p = 0.439$ $\beta = 0.091$	$p = 0.043$; $\beta = -0.257$

Table 6. Significance and probability of Type II error (β) for each trial type (dependent variable; accuracy) in the categorization task by predictors of learning.

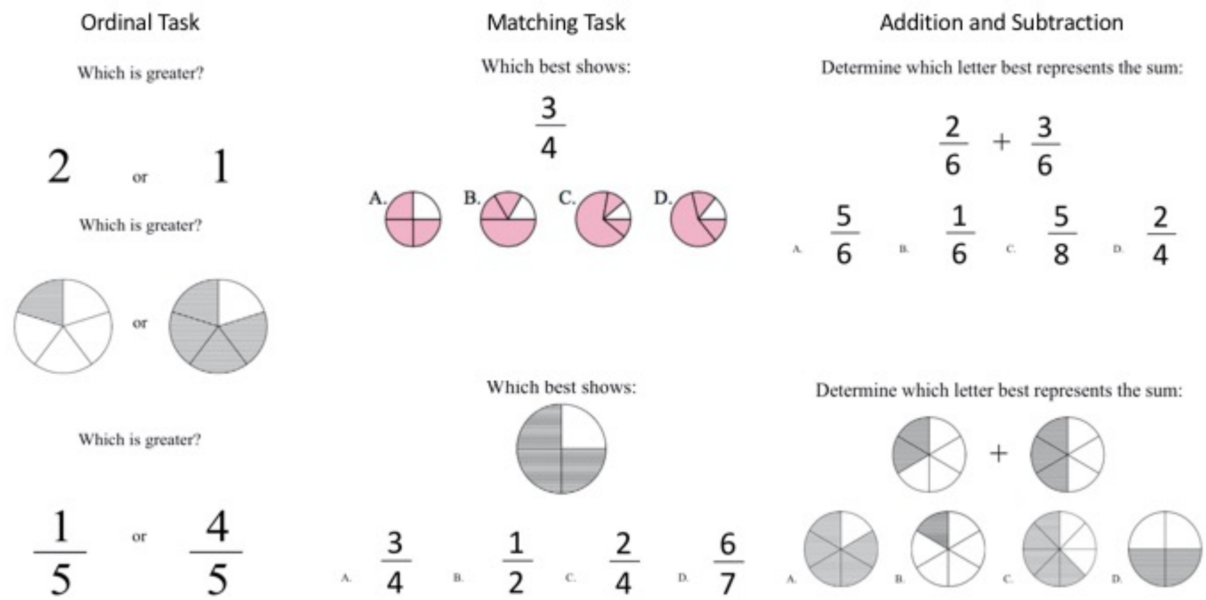

Figure 1. Examples of problem sets in the fraction battery.

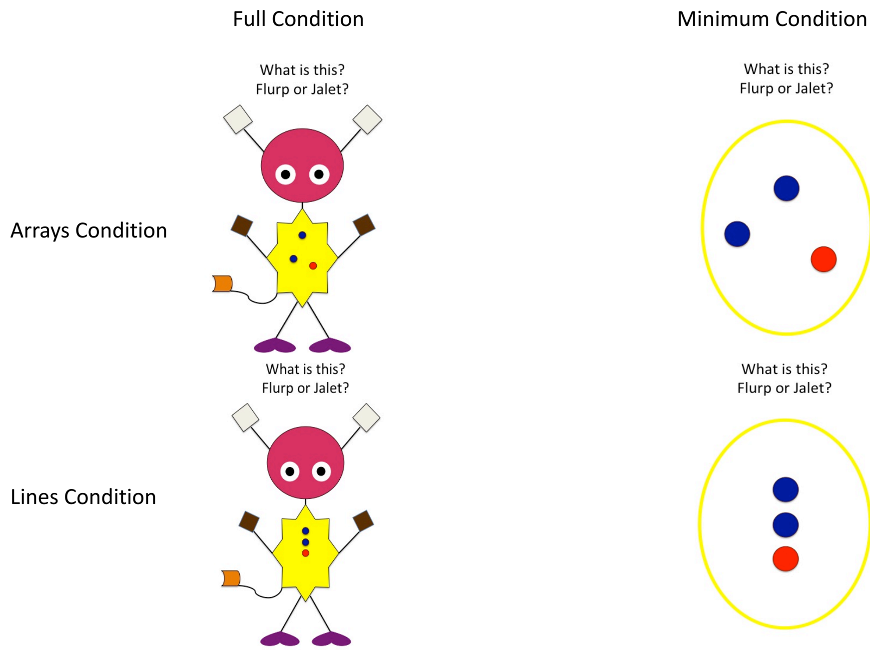

Figure 2. Example of stimuli used across the four conditions in our categorization task.

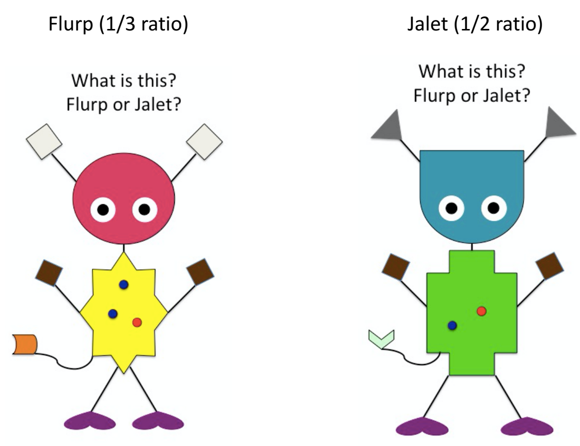

Figure 3. Stimuli examples of a Flurp, with buttons shown in a 1/3 ratio, and a Jalet, with buttons in a 1/2 ratio.

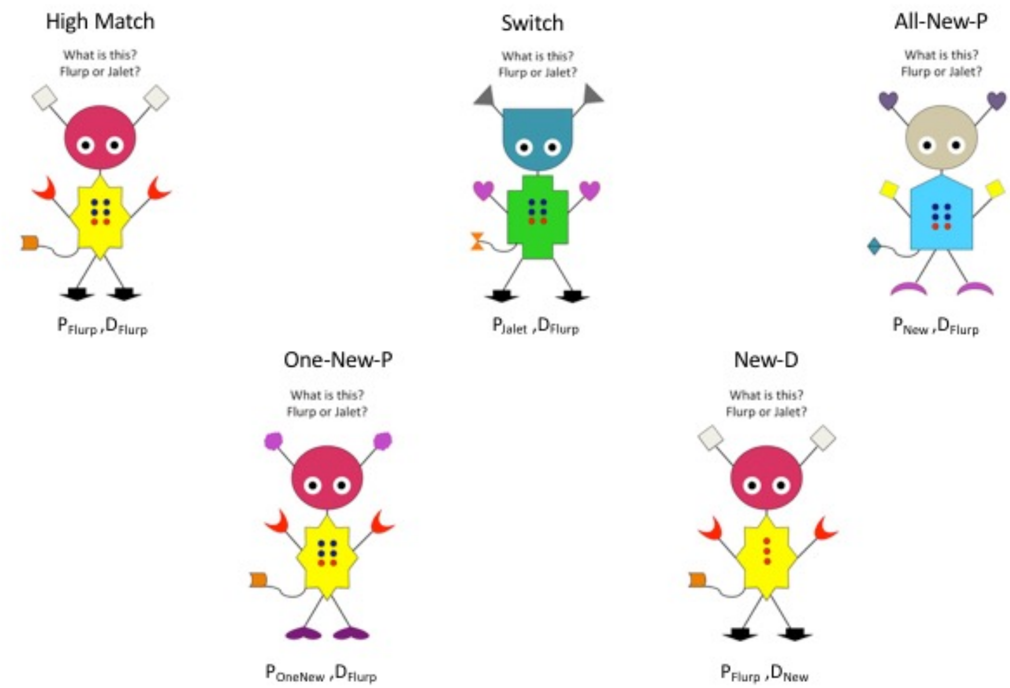

Figure 4. Examples of five trial types used in our categorization task.

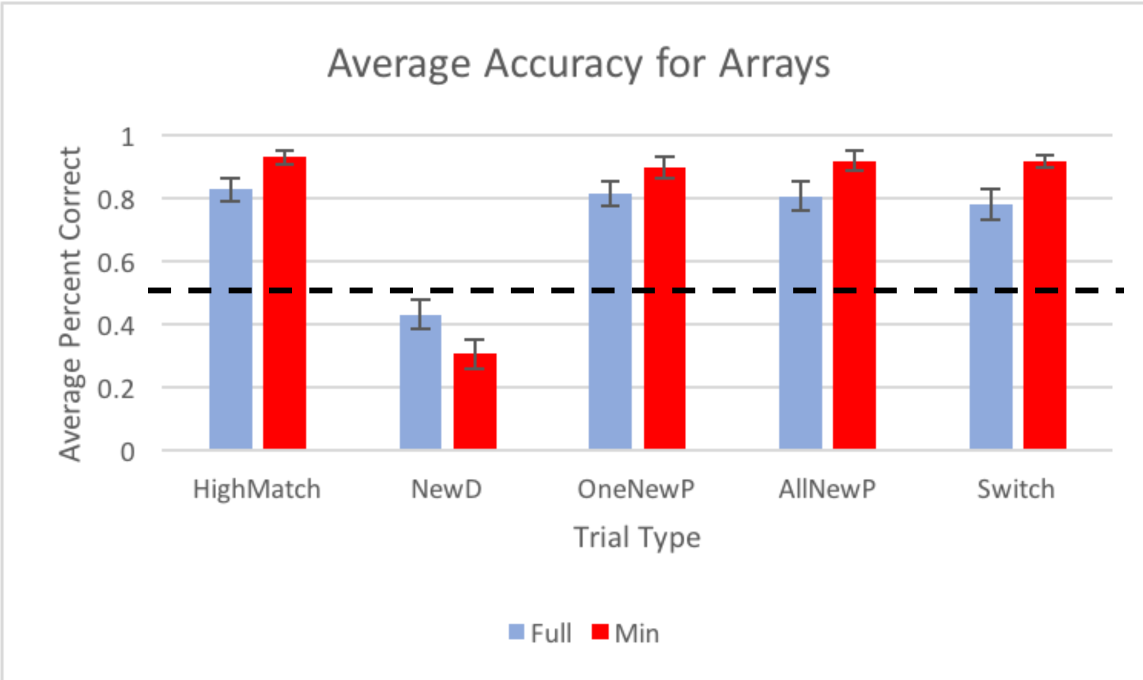

Figure 5. Results of a 5 (Trial Type) x 2 (Full vs. Minimum) x 2 (Lines vs. Arrays) ANOVA demonstrating a main effect of Full vs. Minimum within the Arrays condition, indicating adults were more accurate within the Minimum condition compared to the Full condition.

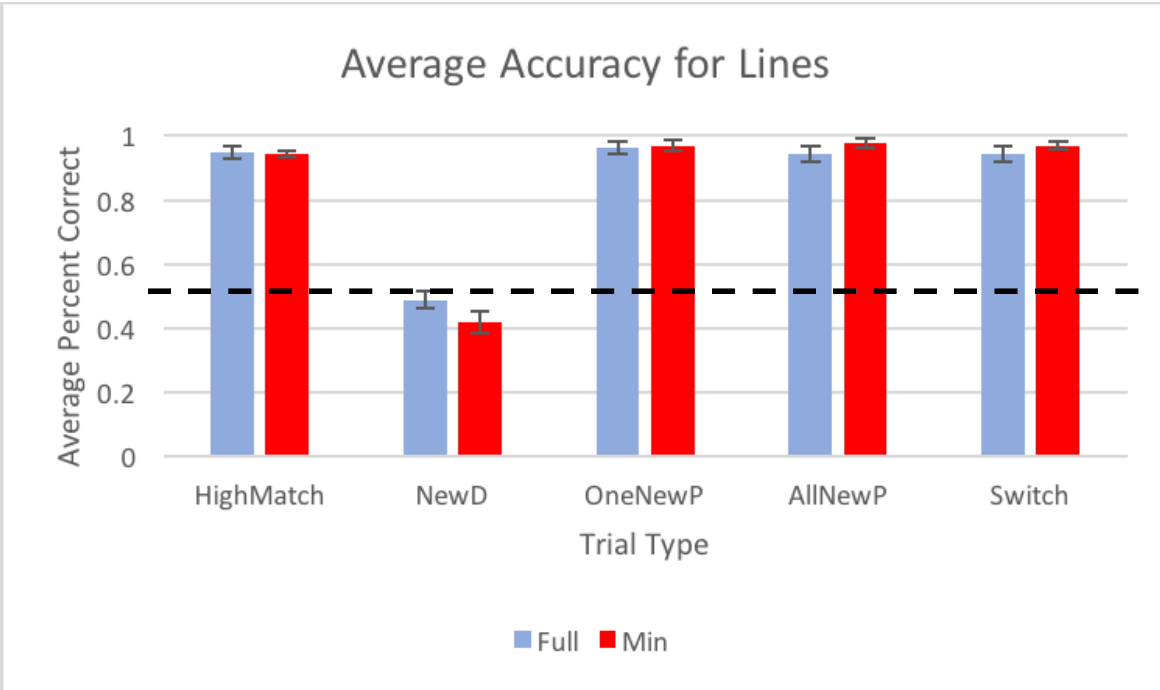

Figure 6. Results of a 5 (Trial Type) x 2 (Full vs. Minimum) x 2 (Lines vs. Arrays) ANOVA demonstrating accuracy in the Lines condition was at-ceiling and a significant main effect of Full vs. Minimum was not discovered.

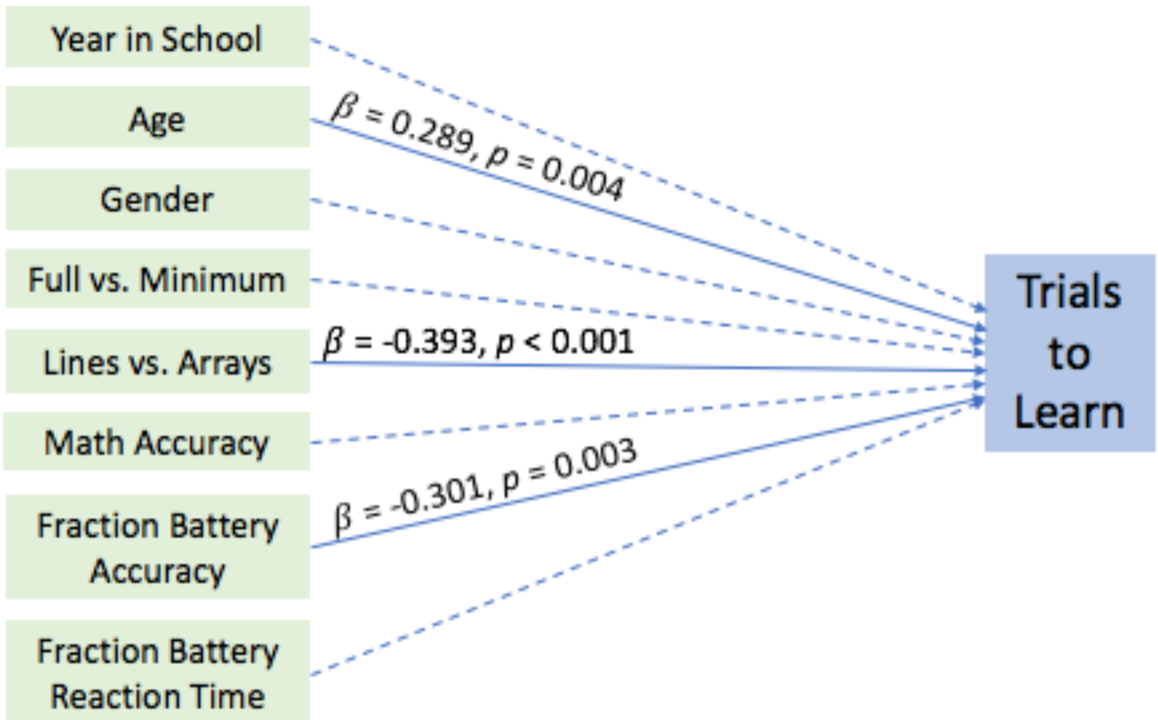

Figure 7. Predictors of learning for Trials to Learn within the training phase of the categorization task. Dotted lines indicate the predictor was nonsignificant; $p > 0.1$; Model: $R^2 = 0.469$, $p < 0.001$.

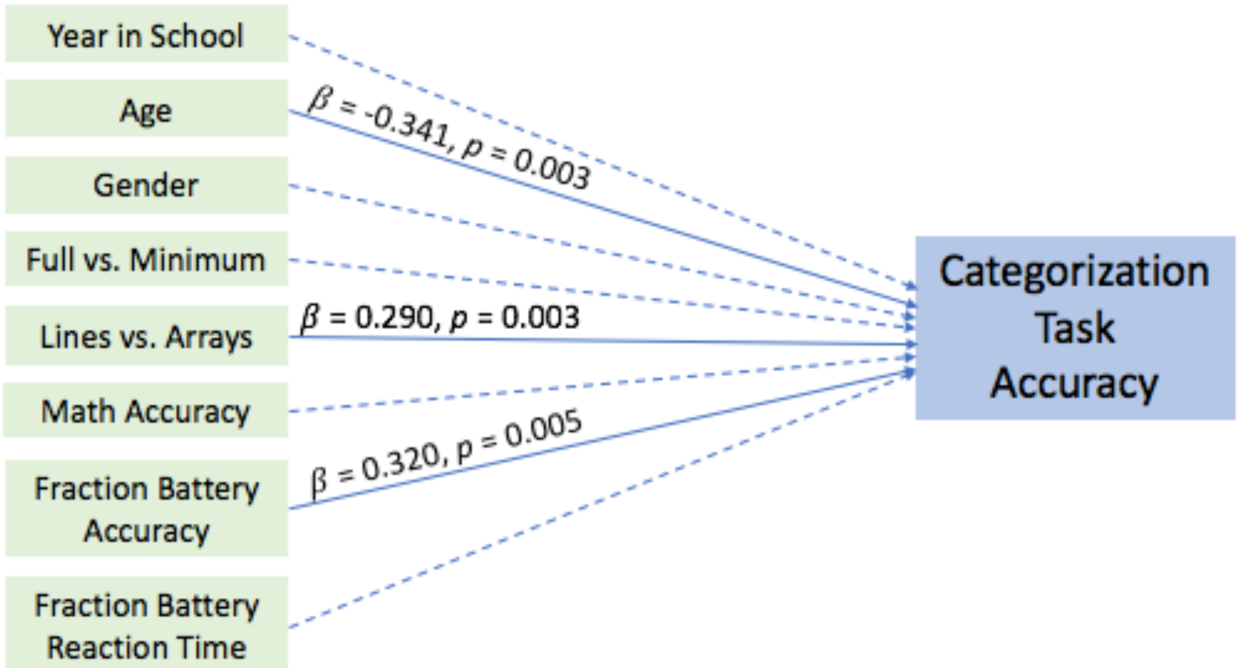

Figure 8. Predictors of learning for Test Trial Accuracy within the categorization task. Dotted lines indicate the predictor was nonsignificant; $p > 0.1$; Model: $R^2 = 0.336$, $p < 0.001$.

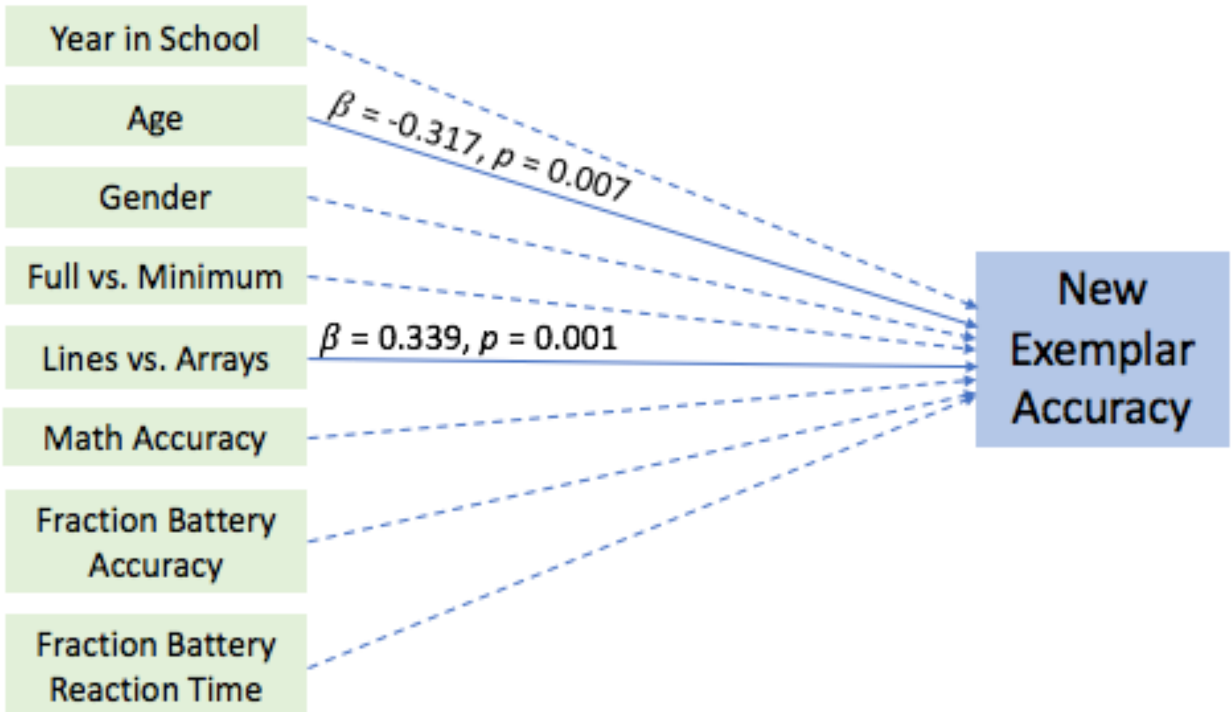

Figure 9. Predictors of learning for New Exemplar (3/9 or 5/10) Accuracy within the categorization task. Dotted lines indicate the predictor was nonsignificant; $p > 0.1$; Model: $R^2 = 0.269$, $p = 0.001$.

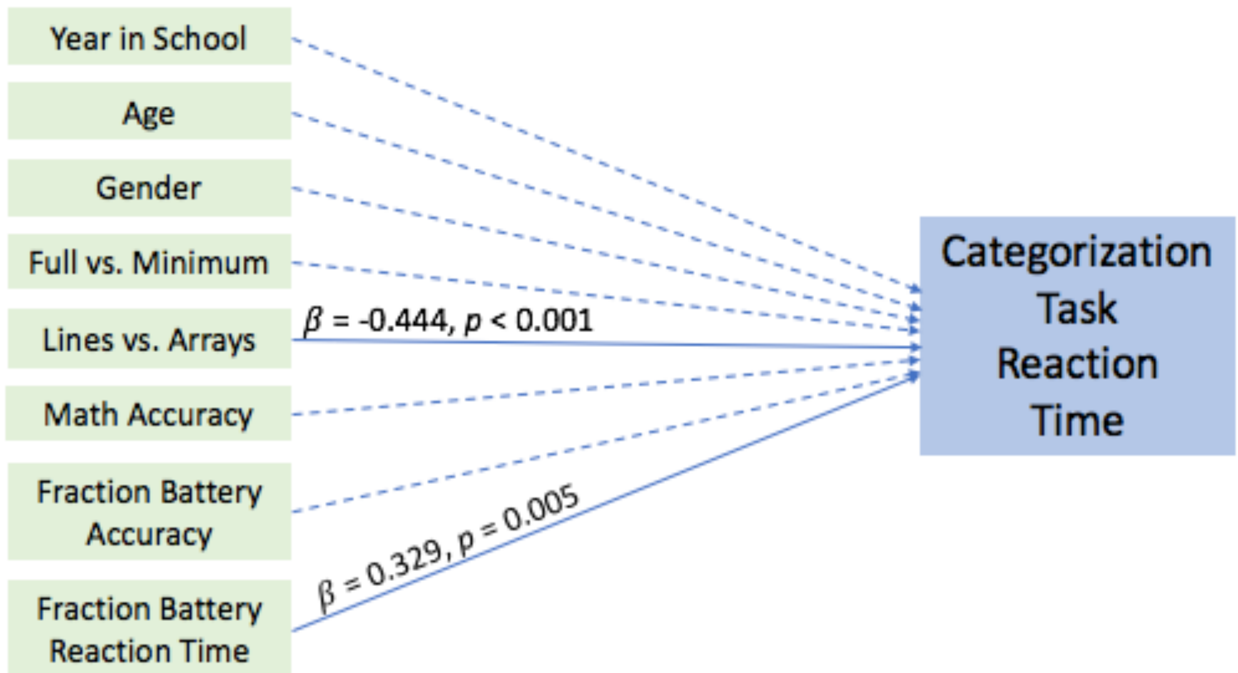

Figure 10. Predictors of learning for Test Trial Reaction Time within the categorization task. Dotted lines indicate the predictor was nonsignificant; $p > 0.1$; Model: $R^2 = 0.313$, $p < 0.001$.